

# MPEG-2

— high-compression technologies for HDTV

## Внимание!

- Данный перевод **НЕ** претендует на аутентичность и может содержать отдельные неточности.
- Оригинал этого документа находится по адресу: <http://www.ebu.ch>

# MPEG-2

— технологии с высокой степенью компрессии для HDTV

Masaaki Kurozumi, Yukihiro Nishida и Eisuke Nakasu  
NHK STRL

Стандарты кодирования цифрового видео обеспечивают гибкость в методах кодирования и допускают усовершенствование эффективности кодирования, сохраняя соответствие стандарту, даже по прошествии времени. Стандарт кодирования видео MPEG-2 [1] использует адаптивную схему кодирования DCT с прогнозируемой компенсацией движения. Объем служебной информации, включая коды векторов движения и режимы кодирования, часто бывает велик для критических последовательностей HDTV на низкой скорости передачи.

Новый метод кодирования NHK [2] – соответствующий MPEG-2 Main Profile – значительно снижает объем служебной информации и делает цифровые услуги HDTV возможными при низких скоростях передачи, сохраняя совместимость с традиционными вещательными приемниками.

В Японии цифровое наземное телевидение (ДТТ) начато в декабре 2003 г., вслед за внедрением цифрового спутникового телевидения в декабре 2000 г. Сегодня высококачественное HDTV охватывает большинство цифровых вещателей страны. Более 7 млн. семей установили цифровые приемники HD. Все службы HDTV используют формат 1080/60i и видео кодирование MPEG-2 Main-Profile, а скорость сейчас около 20 Mbit/s для цифрового спутникового вещания и около 14 Mbit/s для наземного. Службы HDTV связаны с услугами вещания мультимедийных данных, включая новости, прогноз погоды, игры и т.д. Данные HDTV используют большую часть пропускной способности.

Одна из основных характеристик цифрового вещания – его *расширяемость*, благодаря которой можно вводить новые виды услуг, сохраняя существующие. Вещатели хотят обогатить услуги мультимедийного вещания или вводить дополнительные ТВ программы или новые типы услуг. Для этого придется существенно снизить скорость передачи битов HDTV без ухудшения качества HDTV и без предъявления новых требований к существующим приемникам.

## Схема кодирования MPEG-2 для чересстрочного видео

Видео стандарт MPEG-2 – общая схема кодирования, применяемая к различным видео форматам с разным разрешением, частотой кадров и структурой развертки. MPEG-2 также предлагает эффективные инструменты, адаптивные к контенту, для кодирования чересстрочных видеосигналов.

## Структура изображения и режим кодирования

В Таблице 1 показаны элементы синтаксиса видео MPEG-2 для прогнозирования и типов кодирования DCT (дискретно-косинусное преобразование – ред.). Выбор двух типов структуры изображения (PS), кадр и поле, выбирается на уровне изображения. Флажок `frame_pred_frame_dct` (FPFD) имеется только для кадровой структуры. Комбинации PS и FPFD, которые можно выбрать на уровне изображения, следующие:

- PS = Изображение кадра, FPFD = '0'
- PS = Изображение кадра, FPFD = '1'
- PS = Изображение поля, FPFD = '0'

Таблица 1  
Элементы синтаксиса типа прогнозирования и типов DCT в видео MPEG-2

Уровень изображения		Уровень макроблоков			
Picture_structure (PS)	frame_pred_frame_dct (FPFD)	macroblock_modes		Тип прогнозирования	Тип DCT
Изображение кадра	'0'	frame_motion_type (2-бит)	dct_type (1-бит)	Полевой	Кадровый DCT или Полевой DCT
				Кадровый	
				Dual-prime	
Изображение кадра	'1'	–	–	Кадровый	Кадровый DCT
Изображение поля	'0'	field_motion_type (2-бит)	–	Полевой	Полевой DCT
				16x8 MC	
				Dual-prime	

В кадровой структуре изображения прогнозирование и DCT макроблока выбираются адаптивно в соответствии с `frame_motion_type` и `dct_type` в `macroblock_modes`. Тип прогнозирования идентифицируется с `frame_motion_type` (2-бит) или `field_motion_type` (2-бит), а тип кодирования DCT идентифицируется с `dct_type` (1-бит). Флажок FPFD используется для сокращения служебной информации для прогрессивных последовательностей. Когда FPFD = '1', режим кодирования макроблока ограничен комбинацией прогнозирования кадра и DCT кадра, и `macroblock_modes` не передается. В полевой структуре изображения `dct_type` не передается, т.к. тип DCT ограничен только полями.

Традиционные кодеры обычно работают с постоянной комбинацией PS = frame и FPFD = '0' (желтые строки в Таблице 1) для кодирования чересстрочных последовательностей. В этом случае 1 бит для изображений I и 3 бита для изображений P и B должны передаваться для идентификации прогнозирования и типов DCT в макроблоке. Объем такой информации для HDTV составляет до 0.7 Mbit/s – что становится существенным при низкой скорости передачи.

Кадровое кодирование изображения поддерживает и кадровое, и полевое прогнозирование, как показано на Рис. 1. Кадровое прогнозирование хорошо работает с неподвижными или очень медленными изображениями, а полевое показывает лучшие характеристики для изображений с быстрым движением. Полевое прогнозирование требует два вектора движения на макроблок и потребляет больше кодов вектора движения по сравнению с кадровым прогнозированием.

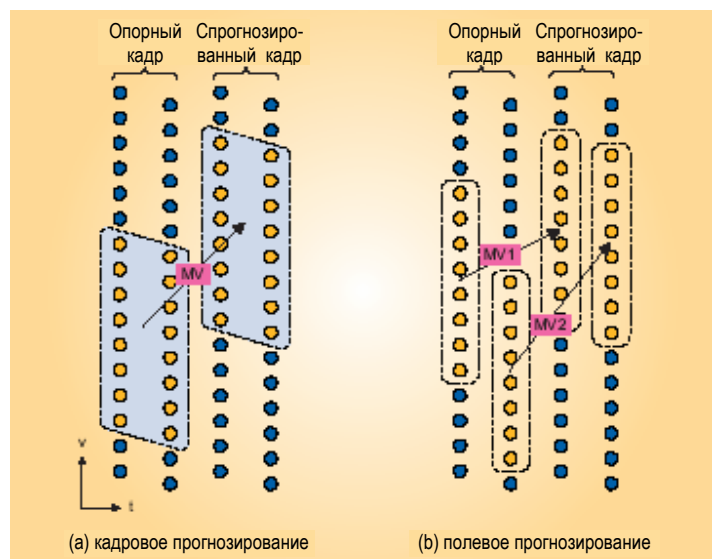


Рис. 1  
Адаптивное прогнозирование кадровых изображений

На Рис. 2 показано два DCT типа кадровых изображений. Кадровый и полевой режимы DCT адаптируются для каждого макроблока в соответствии с текстурой и движением сцены.

### Проблемы в кодировании с низкой скоростью передачи

Кодированный поток битов видео MPEG-2 классифицируется синтаксически режимом кодирования, векторами движения и коэффициентами DCT.

На Рис. 3 и 4 показана статистика потоков битов для информации режима кодирования (Header), векторов движения (MV) и коэффициентов DCT (DCT) при использовании традиционного кодера MPEG-2, использующего кадровую структуру изображения (PS = frame) и FPFD = '0'. Поскольку коды векторов движения составляют большинство битов кодирования для последовательностей HDTV с быстрым движением независимо от запланированной скорости потока, то коэффициентам DCT нельзя присвоить адекватные биты кодирования, что ведет к ухудшению качества изображения на низкой скорости.

### Проблемы в кодировании компонентов цветности в чересстрочном видео

Есть две главные причины деградации цвета в быстро движущихся изображениях чересстрочного видео:

- неверное предсказание выборки цветности, и
- неадаптивный режим DCT для компонентов цветности.

Большинство кодеров используют для оценки движения только выборки яркости из-за их простой реализации. Однако это иногда вызывает заметную деградацию цвета. Кроме того, в кадровом предсказании чересстрочного видео 4:2:0 предсказание выборки

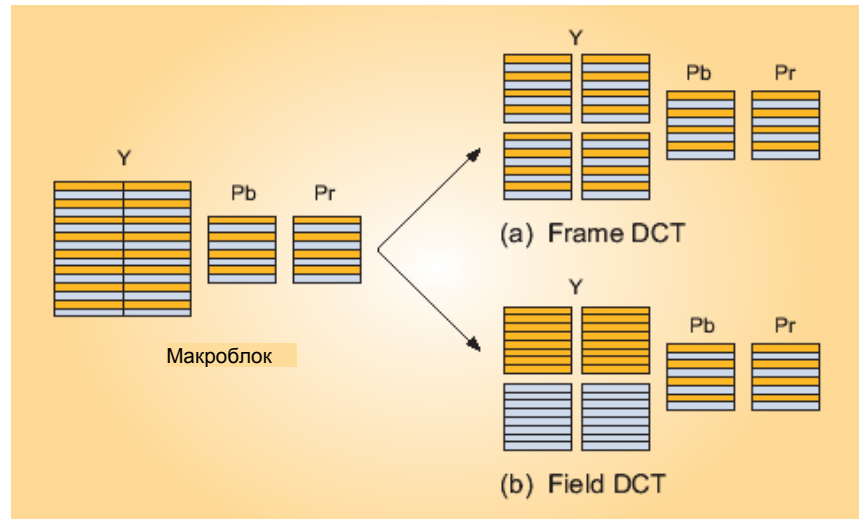


Рис. 2  
Адаптивное DCT кодирование кадровых изображений

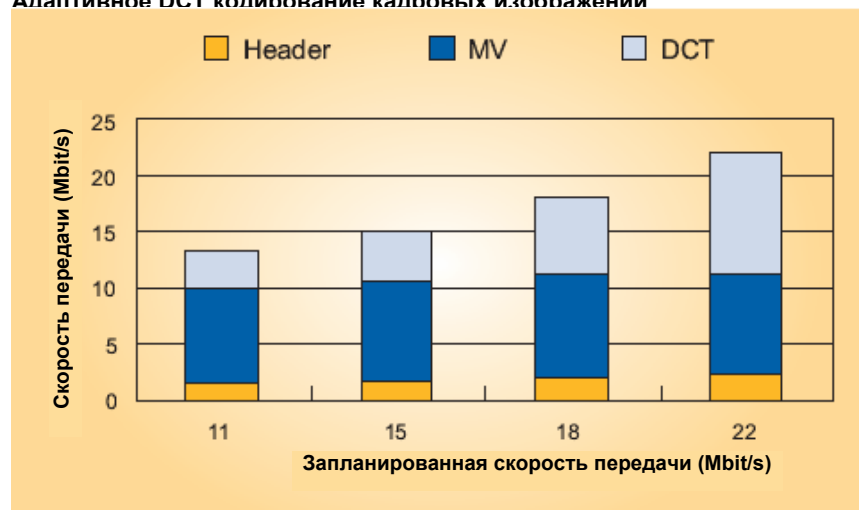


Рис. 3  
Общая статистика потока битов в традиционном методе ("Whale Show", MP@HL)

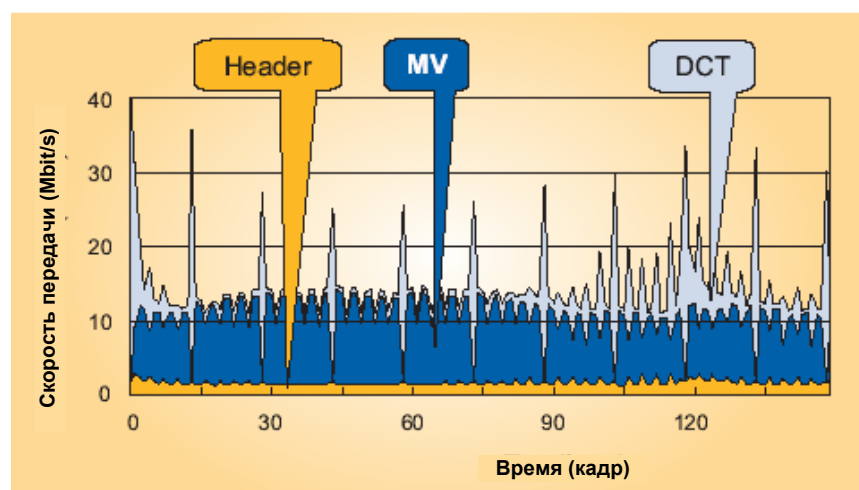


Рис. 4  
Покадровая статистика потока битов в традиционном методе ("Whale Show", MP@HL 15 Mbit/s)

цветности в двух полях может использовать не паритетные отсчеты и не оптимальные временные сдвиги, как показано на *Рис. 5*. Такое неэффективное предсказание увеличит ошибки для компонентов цветности и вызовет деградацию цветности.

Что касается режима DCT, как показано на *Рис. 2*, хотя для компонентов яркости есть и кадровый, и полевой режимы DCT, для компонентов цветности доступен только кадровый DCT. Поэтому кадровый DCT используется для кодирования компонентов цветности даже для быстро движущихся изображений, для которых больше подходит полевой режим DCT. Это вызывает высокие коэффициенты в вертикальных высокочастотных компонентах для чересстрочных изображений и значительную деградацию цветности.

## Новый метод кодирования для решения проблем

Для решения проблем кодирования с низкой скоростью передачи битов необходимо сократить служебную информацию, включая коды векторов движения и адаптивные режимы кодирования. Проблемы кодирования компонентов цветности можно решить адаптивно к контенту, используя полевое изображение и компоненты цветности для оценки движения. С учетом этого наш новый метод кодирования использует две схемы:

- 1) адаптивный выбор структуры изображения с принудительным режимом кодирования, и
- 2) оптимизированный выбор векторов движения и режимов кодирования макроблоков в синтаксисе MPEG-2.

### Адаптивный выбор структуры изображения

Структура изображения адаптивно выбирается как кадровая или полевая в соответствии с содержанием изображения. Объем кодовых битов увеличивается, когда в коэффициенты DCT включаются высокочастотные элементы. Предлагаемый метод определяет адекватную структуру изображения для каждой картинки, сравнивая оценки на основе энергии частот кадров.

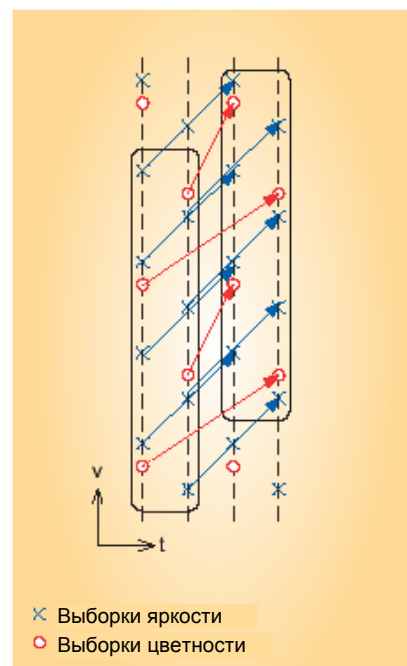
Следующие комбинации PS и FPFD (зеленые строки в *Таблице 1*) используются для сокращения объема информации о режиме для предсказания и типов DCT в кадровом изображении:

- PS = Кадровое изображение, FPFD = '1'
- PS = Полевое изображение, FPFD = '0'

В случае PS = frame и FPFD = '1' изображение состоит только из кадровых блоков DCT, а в случае PS = field и FPFD = '0' – только из полевых блоков DCT.

### Оптимизированный выбор векторов движения и режимов кодирования макроблоков

Новый метод вводит правило выбора с целью максимизации эффективности кодирования, которое является функцией ошибки предсказания и количества битов вектора движения и режима макроблока. В каждом макроблоке вектор движения и режим макроблока выбираются так, чтобы ошибка предсказания была меньше, а присваиваемое количество битов для коэффициентов DCT – больше. Для каждого возможного «межмакроблочного» (intermacroblock) режима отношение присваиваемого количества битов для коэффициентов DCT к средней абсолютной ошибке предсказания с компенсацией движения оценивается во время поиска вектора движения, и вектор движения, дающий максимальное значение, выбирается для каждого межмакроблочного режима. Этот режим представляет комбинации направления предсказания (вперед, назад или двунаправленно) и тип предсказания (кадровое, полевое, 16x8MC и dual-prime). Присваиваемое количество битов для коэффициентов DCT вычисляется путем вычитания количества битов режима макроблока и вектора движения из запланированного количества битов для макроблока.



**Рис. 5**  
Пример опорных выборов в кадровом предсказании

Режим макроблока, дающий максимальную эффективность кодирования, выбирается из всех возможных режимов макроблока. Во «внутримакроблочном» (intra macroblocks) режиме, поскольку значение квантования внутреннего коэффициента DC постоянно внутри изображения, эффективность кодирования вычисляется как отношение присваиваемого количества битов для коэффициентов AC к дисперсии входящих пикселей в блоке. В межмакроблочном режиме эффективность кодирования вычисляется как отношение присваиваемого количества битов для всех коэффициентов DCT к среднеквадратической ошибке предсказания с компенсацией движения.

Вышеупомянутые процессы для определения оптимального набора векторов движения и режима макроблоков используют компоненты яркости и цветности.

## Результаты экспериментов

Эффективность нового метода оценивалась путем моделирования кодирования. Также было проведено сравнение с традиционным кодером MPEG-2, использующим PS = frame и FPFD = '0'. В Таблице 2 показаны общие параметры кодирования. Методы использовали одно и то же управление скоростью передачи. Использовалось одинаковое запланированное количество битов для макроблока в изображении на основе управления скоростью передачи методом TM5 [3].

Рис. 6 и 7 иллюстрируют статистику потока битов нового метода. Если сравнить Рис. 6 и 7 с Рис. 3 и 4, видно, что новый метод значительно снижает объем кодов векторов движения и увеличивает число битов, присваиваемых коэффициентам DCT. Рис. 8 иллюстрирует улучшение PSNR в новом методе по сравнению с PSNR традиционного кодера. Рис. 9 сравнивает PSNR нового метода и традиционного кодера по трем тестовым материалам HDTV, закодированным на скорости 11, 15, 18 и 22 Mbit/s. Улучшение PSNR при 11 и 15 Mbit/s было от 1 до 4 dB. PSNR нового метода при 15 Mbit/s и 11 Mbit/s было сопоставимо с PSNR традиционного кодера при скорости 18 – 22 Mbit/s и 15 – 18 Mbit/s.

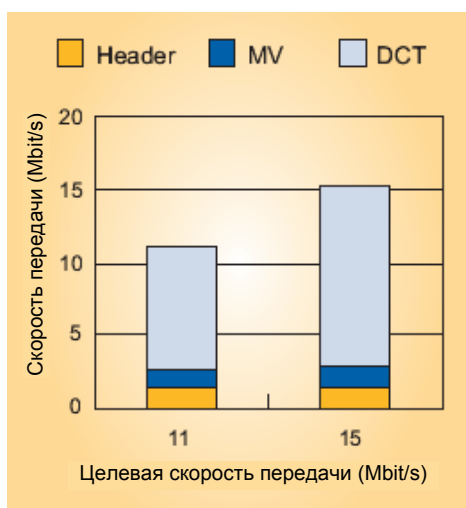


Рис. 6  
Общая статистика потока битов нового метода ("Whale Show", MP@HL)

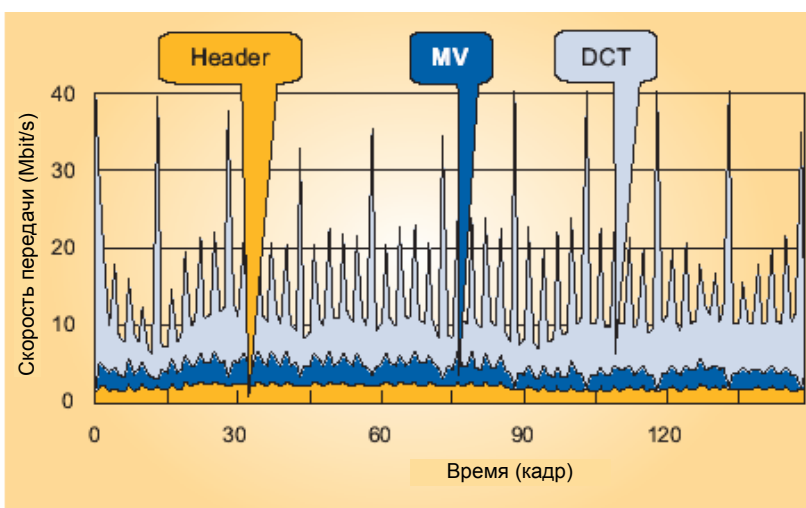


Рис. 7  
Покадровая статистика потока битов нового метода ("Whale Show", MP@HL 15 Mbit/s)

Таблица 2  
Условия тестирования

Формат изображения	1920 × 1080 /59.94i
Формат цветности	4:2:0
Структура GoP	N=15, M=3
Точность внутри DC	8 бит
Тип шкалы квантизатора	Нелинейный
Область поиска векторов движения (по горизонтали × по вертикали)	±63.5 × ±31.5 на кадровый интервал
Тестовые последовательности	Тестовые материалы HDTV ITU-R; "European Market (Euro)", "Whale Show (Whale)" и "Green Leaves (Green)"

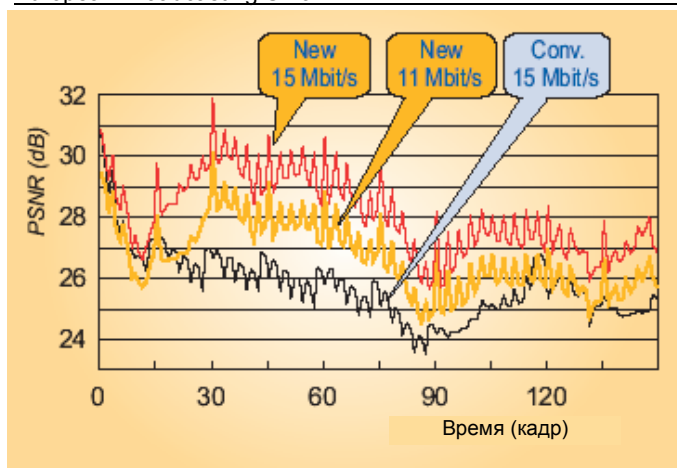


Рис. 8  
Улучшение PSNR в новом методе (“Whale Show”)

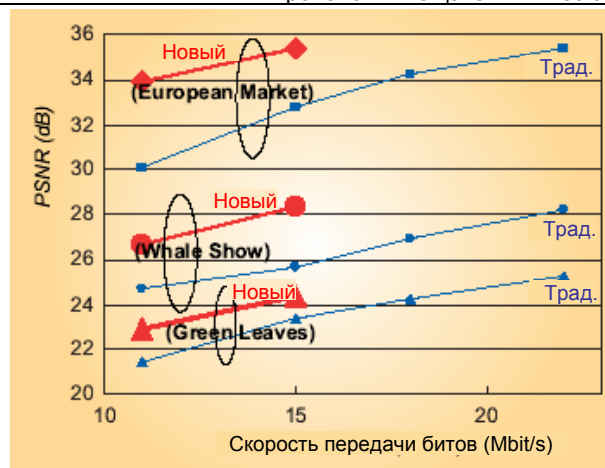


Рис. 9  
Производительность PSNR в новом методе

Рис. 10 (на следующей странице) показывает вырезанные участки изображений HDTV, закодированные при 15 Mbit/s традиционным кодером и новым методом. Новый метод уменьшает искажение блоков и значительно повышает субъективное качество изображения. Он сохраняет пространственные детали во всех сценах, а повышение качества изображения особенно заметно в быстро движущихся сценах.

## В заключение

Стандарт видео кодирования MPEG-2 обеспечивает несколько алгоритмов кодирования, адаптивных к контенту. Если видео кодер достаточно использует алгоритмы видео кодирования согласно стандарту, можно повысить эффективность кодирования, сохраняя совместимость со старыми приемниками. Новый указанный здесь метод кодирования эффективно сокращает служебную информацию и повышает качество изображения. Ключевые технологии этого метода включают **адаптивный выбор структуры изображения** и **оптимизированный выбор режима кодирования**, оба соответствующие MPEG-2 Main Profile.

Несмотря на появление и ожидаемое внедрение в новых цифровых медиа, включая вещание, новых стандартов видео кодирования, например, MPEG-4 AVC/H.264, MPEG-2 еще проживет некоторое время – особенно в ситуациях, где он уже используется. Новый метод кодирования для MPEG-2 с высокой степенью компрессии будет хорошим вариантом для вещателей и производителей, желающих повысить качество изображения или сократить емкость канала.

## Сокращения

<b>DCT</b>	Discrete Cosine Transform Дискретное косинусное преобразование	<b>MP@HL</b>	(MPEG-2) Main Profile at High Level
<b>DTT</b>	Digital Terrestrial Television Цифровое наземное телевидение	<b>MV</b>	Motion Vector Вектор движения
<b>FPPD</b>	(MPEG) frame_pred_frame_dct	<b>PS</b>	(MPEG) Picture Structure Структура изображения
<b>HDTV</b>	High-Definition Television Телевидение высокой четкости	<b>PSNR</b>	Peak Signal-to-Noise Ratio Пиковое отношение сигнал-шум

## Ссылки

- [1] ISO/IEC 13818-2:2000 | ITU-T Rec. H.262 (2000): **Information technology – Generic coding of moving pictures and associated audio information: Video**  
<http://www.itu.int/ITU-T/publications/recs.html>
- [2] M. Kurozumi et al.: **Lower bitrate HDTV MPEG-2 coding by adaptive selection of picture structure**  
Proc. of PCS 2003, pp. 317 to 320 (2003) [http://www-doc.irisa.fr/cgi-bin/bibvisu?OCLC\\_W3=34265](http://www-doc.irisa.fr/cgi-bin/bibvisu?OCLC_W3=34265)
- [3] ISO/IEC JTC1/SC29/WG11/N0400 (1993): **“Test Model 5”**



(a) Традиционный кодировщик MPEG-2



(b) Новый метод

**Рис. 10**  
Кодированное изображение последовательности “Whale Show” (MP@HL 15Mbit/s)  
Примечание: показанный размер изображения 640x480 пикселей; вырезано из 1920x1080 пикселей



**Masaaki Kurozumi** получил степень BE по электронике в университете Doshisha, Япония, в 1993 г. и степень ME по обработке информации в Институте науки и технологии Nara Institute в 1995 г. В том же году поступил в NHK (японскую вещательную корпорацию) и работал инженером вещания на станции в Осаке. С 1997 г. работает в научно-технических лабораториях NHK и участвует в исследованиях кодирования изображения.

**Yukihiro Nishida** – заместитель директора по визуальным информационным технологиям в научно-исследовательских лабораториях NHK. После окончания университета University, Япония, со степенью магистра электротехники, работает в NHK (японской вещательной

корпорации) и участвует в исследованиях и развитии видео кодирования, измерения качества, цифрового вещания, HDTV и т.д. Часто участвует в стандартизации в области вещания и является вице-председателем рабочей группы 6A ITU-R. Участвует в проектной группе EBU V/TQE (Television Quality Evolution). В 1992 г. провел 6 месяцев в исследовательском центре RAI (Италия), работая над технологией видео кодирования.



**Eisuke Nakasu** получил степени BE и ME по электротехнике в университете Keio, Япония, в 1980 и 1982 гг. соответственно. Поступил в NHK (японскую вещательную корпорацию) в 1982 г. и отвечает за исследование и развитие цифровых вещательных систем, видео кодирование и оценку качества. Является ведущим инженером по визуальным информационным технологиям в научно-исследовательских лабораториях NHK. Участвовал в деятельности по стандартизации MPEG-2 и внес вклад в развитие и стандартизацию систем цифрового спутникового и наземного вещания в Японии. Сейчас является председателем исследовательской группы по методам оценки качества в ARIB, Япония.